Optimization of Sampling Times for PK/PD Models:

Approximation of Elemental Fisher Information Matrix

Valerii Fedorov and Sergei Leonov
(*GSK Research Statistics Unit*)

PODE 2010, Berlin

June 11, 2010

Comparison of software tools

PODE 2007, 2009 meetings:
- PFIM (developed in INSERM, Université Paris 7, France);
- PkStaMp (GlaxoSmithKline, Collegeville, U.S.A.);
- PopDes (CAPKR, University of Manchester, UK);
- PopED (Uppsala University, Sweden) and
- WinPOPT (University of Otago, New Zealand).

Key for model-based optimal designs: Fisher information matrix of a properly defined single observational unit

This presentation:

- Some results of comparison at PODE 2009
- Certain options of calculating/approximating FIM

PODE 2009 comparison

Goal: compare results (information matrix) for a particular model and particular sequence of sampling times

Model: one-compartment, 1st order absorption, single dose $D = 70$ mg at $t = 0$

$$f(x, \boldsymbol{\gamma}) = \frac{Dk_a}{V(k_a - k_e)} \left(e^{-k_e x} - e^{-k_a x}\right). \qquad (1)$$

Response parameters $\boldsymbol{\gamma} = (k_a, CL, V),\ k_e = CL/V$

Individual parameters

$$\boldsymbol{\gamma}_i = \boldsymbol{\gamma}^0 e^{\eta_i},\ \eta_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}), \qquad (2)$$

$\boldsymbol{\gamma}^0 = (1, 0.15, 8),\ \boldsymbol{\Omega} = diag(0.6, 0.07, 0.02)$

PODE 2009 comparison (cont.)

Measurements:

$$y_{ij} = f(\boldsymbol{\gamma}_i, x_{ij}) \; [1 + \varepsilon_{M,ij}] \; + \; \varepsilon_{A,ij}, \qquad (3)$$

$$\{x_{ij}\} \equiv \mathbf{x} = (0.5, 1, 2, 6, 24, 36, 72, 120) \text{ hours}$$

$$\varepsilon_{A,ij} \sim \mathcal{N}(0, \sigma_A^2), \; \varepsilon_{M,ij} \sim \mathcal{N}(0, \sigma_M^2),$$
$$\sigma_A^2 = 0, \sigma_M^2 = 0.01; \quad i = 1, \ldots, 32; \; j = 1, \ldots, 8$$

Combined parameter

$$\boldsymbol{\theta} = (k_a^0, \; CL^0, \; V^0; \;\; \omega_{k_a}^2, \; \omega_{CL}^2, \; \omega_V^2; \;\; \sigma_M^2)$$

Individual information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$

$\mathbf{x} = (x_1, x_2, \ldots, x_k)$ - sequence of sampling times

$\mathbf{Y} = [y(x_1), \ldots, y(x_k)]^T$ - observations

$\mathbf{f} = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{\theta}}(\mathbf{Y}) = [f(x_1, \boldsymbol{\theta}), \ldots, f(x_k, \boldsymbol{\theta})]^T$ (mean)

$\mathbf{S} = \mathbf{S}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{Var}_{\boldsymbol{\theta}}(\mathbf{Y})$ (variance)

Key formula for Gaussian observations $\mathbf{Y} \sim \mathcal{N}(\mathbf{f}, \mathbf{S})$:

$\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta}) = [\mu_{\alpha\beta}(\mathbf{x}, \boldsymbol{\theta})]_{\alpha,\beta=1}^{m}$, Magnus and Neudecker (1988)

$$\mu_{\alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial \mathbf{f}}{\partial \theta_\alpha} \mathbf{S}^{-1} \frac{\partial \mathbf{f}}{\partial \theta_\beta} + \frac{1}{2} \operatorname{tr} \left[\mathbf{S}^{-1} \frac{\partial \mathbf{S}}{\partial \theta_\alpha} \mathbf{S}^{-1} \frac{\partial \mathbf{S}}{\partial \theta_\beta} \right] . \quad (4)$$

Individual information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ (cont.)

Once $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ is calculated (approximated) for any candidate sequence $\mathbf{x}$, the numerical construction of locally optimal designs is easy!

- Define normalized matrix $\mathbf{M}(\xi, \boldsymbol{\theta})$, w/o or with costs,

$$\mathbf{M}(\xi, \boldsymbol{\theta}) = \sum_i p_i \boldsymbol{\mu}(\mathbf{x}_i, \boldsymbol{\theta}), \;\; p_i \in [0, 1], \;\; \xi = \{p_i, \mathbf{x}_i\}$$

- Specify a criterion of optimality Ψ (D-, A-, c- etc.)

- Solve optimization problem $\Psi\left[\mathbf{M}^{-1}(\xi, \boldsymbol{\theta})\right] \rightarrow \min_\xi$ (1st order optimization algorithm in the space of information matrices $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta}), \;\; \mathbf{x} \in \mathbf{X}$ - design region)

Information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ (cont.)

- Formula (4) is exact for normal $\mathbf{Y}$ only
- Need expressions (approximations?) of $\mathbf{f}$ and $\mathbf{S}$
- First-order approximation (Taylor series):

$$\mathbf{E}_{\boldsymbol{\theta}}(\mathbf{Y}) = [f(x_1, \boldsymbol{\gamma}^0), .., f(x_k, \boldsymbol{\gamma}^0)]^T = \mathbf{f}(\mathbf{x}, \boldsymbol{\gamma}^0) \quad (5)$$

$$\mathbf{Var}_{\boldsymbol{\theta}}(\mathbf{Y}) = \mathbf{S}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{F}\ \boldsymbol{\Omega}\ \mathbf{F}^T \ + \ \sigma_A^2 \mathbf{I}_k \ +$$
$$+ \ \sigma_M^2 \ Diag\left[\mathbf{f}(\mathbf{x}, \boldsymbol{\gamma}^0)\mathbf{f}^T(\mathbf{x}, \boldsymbol{\gamma}^0) \ + \ \mathbf{F}\ \boldsymbol{\Omega}\ \mathbf{F}^T\right], \quad (6)$$

$\mathbf{F} = \left[\frac{\partial \mathbf{f}(\mathbf{x}, \boldsymbol{\gamma}^0)}{\partial \gamma_\alpha}\right]$ $-(k \times m_1)$-matrix, $m_1 = dim(\boldsymbol{\gamma}) = 3$

Gagnon and Leonov (2005)

Information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ (cont.)

First round of PODE 2009 comparison, expression for FIM: very similar results for coeff. of variation (CV) except k_a

(i) PFIM, PopED, WinPOPT: $CV(k_a) \simeq \mathbf{13.9}\%$

(ii) PkStaMp, PopDes: $CV(k_a) \simeq \mathbf{4.8}\%$, why such discrepancy??

(iii) Simulations in NONMEM/MONOLIX: $CV(k_a) \simeq \mathbf{12 - 13}\%$, why closer to (a)??

Information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ (cont.)

Matrix $\boldsymbol{\mu}$ in (4): block form, *Retout and Mentré (2003)*

$$\boldsymbol{\mu} = \left\{ \begin{array}{cc} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{array} \right\}, \tag{7}$$

$\mathbf{A} = \mathbf{F}^T\, \mathbf{S}^{-1}\, \mathbf{F} + \frac{1}{2}\,\mathrm{tr}$ (derivatives wrt γ_α)
$\mathbf{C} = \frac{1}{2}\,\mathrm{tr}$ (mixed derivatives wrt γ_α and $[\omega_\beta^2, \sigma_M^2]$)
$\mathbf{B} = \frac{1}{2}\,\mathrm{tr}$ (derivatives wrt $[\omega_\beta^2, \sigma_M^2]$)

PkStaMp in (i): used (5), (6) and FULL matrix $\boldsymbol{\mu}$ in (7)

If (1) Block $\mathbf{C}$ "excluded" $(\mathbf{C} = \mathbf{0})$
(2) Second term in $\mathbf{A}$ (trace) removed
(3) No term $\mathbf{F}\ \boldsymbol{\Omega}\ \mathbf{F}^T$ in square brackets in (6) $\Longrightarrow$
then exact match with (ii)

Which approximation to choose?

Types of approximation

A1. Log-normal distribution in (2)

- *1st-order approximation*, $\mathbf{E}\eta_i = 0,\ \mathbf{Var}(\eta_i) = V \implies$

$$\mathbf{E}_\eta(\theta e^{\eta_i}) \simeq \theta, \qquad \mathbf{Var}_\eta(\theta e^{\eta_i}) \simeq \theta^2 V$$

- *Exact moments*:

$$\mathbf{E}_\eta(\theta e^{\eta_i}) = \theta e^{V/2}, \quad \mathbf{Var}_\eta(\theta e^{\eta_i}) = \theta^2 e^V (e^V - 1).$$

- If $\theta = 1,\ V = 0.6$ (as for k_a), then

$\mathbf{E}_{1st} = 1,\ \mathbf{E}_{exact} = 1.35;\ \mathbf{Var}_{1st} = 0.6, \mathbf{Var}_{exact} = 1.50$

Types of approximation (cont.)

A2. Trace (2nd) term in (4): let

- $k = 1$ (single response) and
- $\mathbf{Var}(y) = S = \sigma^2 f^2$ with known σ^2; cf. (3)

$$\Downarrow$$

$$\boldsymbol{\mu} = \frac{1}{\sigma^2}\,\frac{\mathbf{F}^T\mathbf{F}}{f^2} \; + \; 2\,\frac{\mathbf{F}^T\mathbf{F}}{f^2} \; = \; \left(\frac{1}{\sigma^2} + 2\right)\,\frac{\mathbf{F}^T\mathbf{F}}{f^2} \qquad (8)$$

To examine the effect of missing 2nd term on CV, check

$$\sqrt{\frac{\mu_{\alpha\beta,\ FULL}}{\mu_{\alpha\beta,\ 1st\ term}}} = \sqrt{\frac{2 + 1/\sigma^2}{1/\sigma^2}} = \sqrt{1 + 2\sigma^2} \sim 1 + \sigma^2 \Longrightarrow$$

<u>Inflation coefficient</u> for CV: $1 + \sigma^2,\ \sigma^2 \leq 0.25$

Types of approximation (cont.)

A3. Second-order approximation for mean/variance:
Fedorov, Leonov (2005)

$$\mathbf{E}_{\boldsymbol{\theta}}[f(x,\gamma_i)] \approx f(x,\boldsymbol{\gamma}^0) + \frac{1}{2}\,\mathrm{tr}\left[\mathbf{H}(\boldsymbol{\gamma}^0)\boldsymbol{\Omega}\right], \qquad (9)$$

$$\mathbf{H}(\gamma^0) = \left[\frac{\partial^2 f(x,\boldsymbol{\gamma})}{\partial\gamma_\alpha\,\partial\gamma_\beta}\right]\Bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0} \; etc \quad \Longrightarrow$$

numerically may be rather tedious

Types of approximation (cont.)

A4. Calculation of mean/variance via Monte Carlo:

$$\widehat{f}(x_j) = \widehat{\mathbf{E}}_{\boldsymbol{\theta}}(y_{ij}) = \frac{1}{N}\sum_{i=1}^{N} y_{ij} \ , \qquad (10)$$

$$\widehat{S}(x_j) = \widehat{\mathbf{Var}}_{\boldsymbol{\theta}}(y_{ij}) = \frac{1}{N}\sum_{i=1}^{N}[y_{ij} - \widehat{f}(x_j)]^2 \implies$$

Numerically straightforward: valid option if normal approximation is "reasonable" (??)

Comparison of approximation options

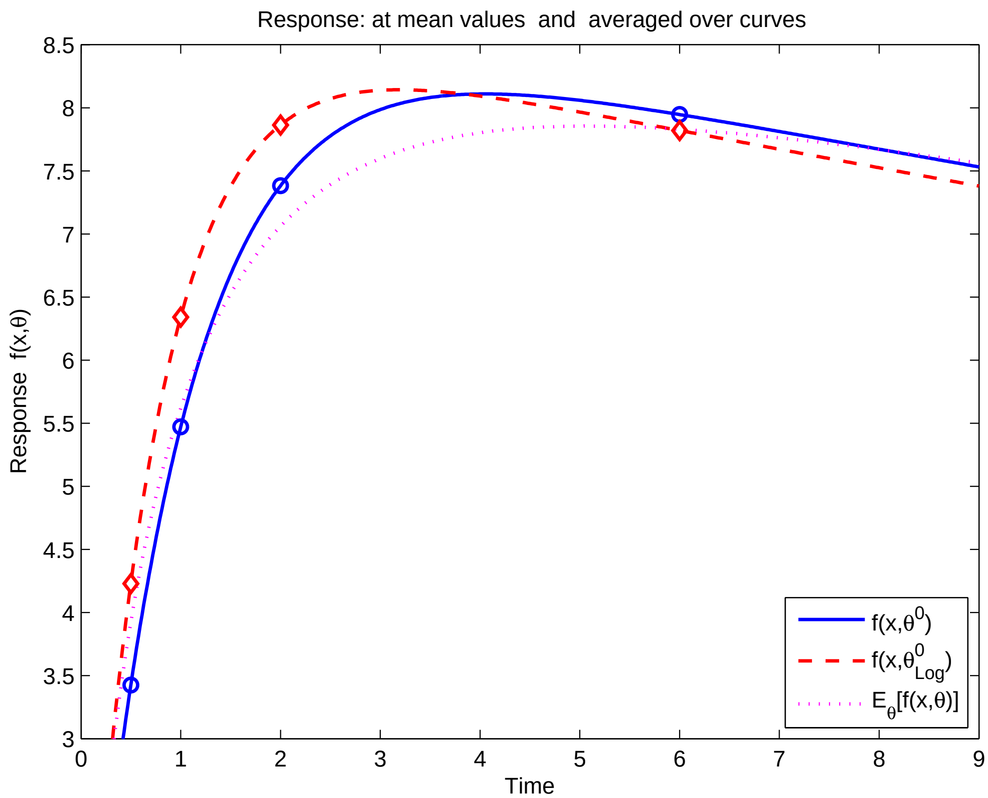

Figure 1: Mean response curves for PODE 2009 example. Solid - 1st order approximation, dashed - computed at mean values of log-normal distribution, dotted - Monte Carlo average as in (10)

For a single response parameter, $m = 1$:

$$\mathbf{E}_\theta[f(x,\theta_i)] \;\approx\; f(x,\theta^0) \;+\; \frac{1}{2}\, f''(x,\theta)\; \mathbf{Var}(\theta)$$

Comparison of approximation options (cont.)

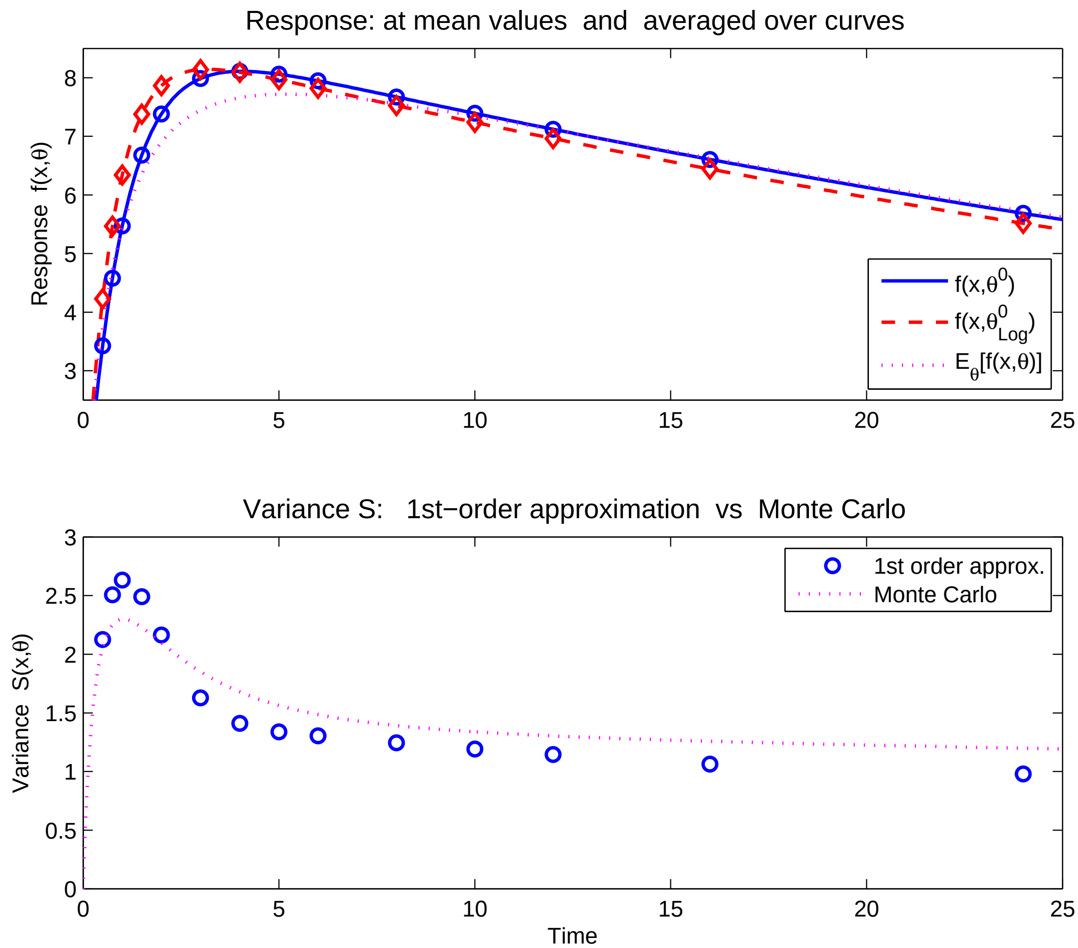

Figure 2: Mean response curves and variance. Legend similar to Fig. 1

References

[1] Atkinson, A.C., and Donev, A. (1992), *Optimum Experimental Design*. Clarendon Press, Oxford.

[2] Fedorov, V.V. (1972), *Theory of Optimal Experiment*. Academic Press, New York.

[3] Fedorov, V.V., and Hackl, P. (1997), *Model-Oriented Design of Experiments*. Springer, NY.

[4] Fedorov, V.V., Leonov, S.L. (2005), Response driven designs in drug development. In: Wong, W.K., Berger, M.P.F. (Eds.), *Applied Optimal Designs*, Wiley, Chichester, pp. 103-136.

[5] Fedorov, V.V., and Leonov, S.L. (2007), Population pharmacokinetic measures, their estimation and selection of sampling times, *J. Biopharm. Stat.*, **17**(5), 919-941.

[6] Gagnon, R., and Leonov, S.L. (2005), Optimal population designs for PK models with serial sampling. *J. Biopharm. Stat.*, **15** (1), 143-163.

[7] Magnus, J.R., and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, NY.

[8] Mentré, F., Duffull, S., Gueorguieva, I., Hooker, A., Leonov, S., Ogungbenro, K., Retout, S. (2007). Software for optimal design in population pharmacokinetics and pharmacodynamics: a comparison In: *Abstracts of the Annual Meeting of the Population Approach Group in Europe (PAGE)*. ISSN 1871-6032, http://www.page-meeting.org/?abstract=1179

[9] Retout, S., and Mentré, F. (2003), Further developments of the Fisher information matrix in nonlinear mixed effects models with evaluation in population pharmacokinetics. *J. Biopharm. Stat.*,**13** (2), 209-227.